

E. H. Melan
R. T. Curtis
J. K. Ho
J. G. Koens
G. A. Snyder

Quality and Reliability Assurance Systems in IBM Semiconductor Manufacturing

Soon after semiconductor manufacturing began it was realized that classical process control techniques were needed for the control of quality, reliability, and yield. The discovery and control of yield, quality, and reliability detractors have been pursued continually by IBM manufacturing engineers ever since, and the resulting evolution of process control techniques has grown into a highly disciplined state. Inferential methods were added later to augment the classic techniques. This paper, in addition to providing a brief overview of semiconductor manufacturing control techniques and placing them into historical perspective, discusses a method of feed-forward control based on statistical distributions which is used in the VLSI FET memory device line. This is followed by a description of a process profile technique which is used in bipolar logic manufacturing. The importance of the system aspects in both techniques is emphasized.

Introduction

It was realized early during the evolution of semiconductor manufacturing that implementation of classical process control techniques in the tradition of Shewhart [1] and Deming [2] was necessary for the control of quality, reliability, and yield of devices. In fact, the invention of the control chart by Shewhart in 1931 was no doubt the most significant development toward quality control of a process step or operation. It provided pictorial results of an operation and compared them to statistically derived limits that indicate degree of control in relation to process capability.

Process-sensitive technologies generally have an exposure to quality and reliability effects that are potentially detrimental in terms of device yields, quality levels, and reliability performance of the product. It is the susceptibility of semiconductor technologies to yield, quality, and reliability detractors that has caused the evolution of process control techniques to the highly disciplined state that exists today within IBM. The density and complexity of VLSI in both logic and memory applications reduces the inspectability of

the devices so that traditional end-of-line testing and feedback control have become insufficient to ensure the necessary quality and reliability levels that are one to two orders of magnitude beyond that of pre-VLSI technologies.

This difficulty has led to the development of inferential methods in controlling product quality and reliability as an accompaniment to classic in-line process control techniques. The complexity of continuous control of very small dimensions and low process impurity levels in device manufacturing has led to the development and implementation of in-line product and process monitors as well as comprehensive data management systems. A feed-forward control approach, based on statistical distributions rather than the traditional specification-limit method, is described in the quality management system used in the manufacture of high-density FET memory devices at IBM's Burlington and Sindelfingen plants. The process profile system described for producing bipolar devices represents the approach used for VLSI manufacturing at IBM's East Fishkill and Essonnes plants.

© Copyright 1982 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

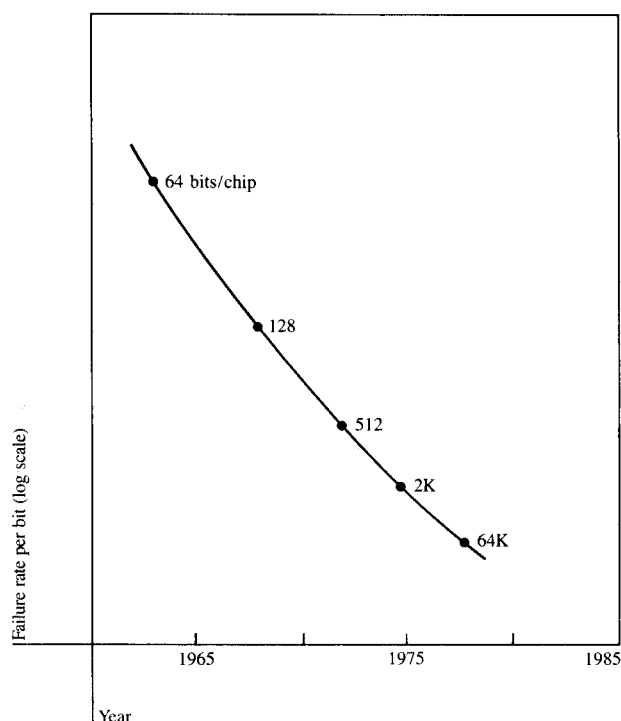


Figure 1 Intrinsic failure rate improvement trends for IBM memory technology. The per-bit percentage failure rate per 1000 hours is shown. The numbers noted on the line refer to the number of bits per chip (K = 1024).

Historical development

SLT technology With the advent of the IBM SLT technology [3] in the middle 1960s, quality control in the manufacturing of discrete transistors and diodes consisted primarily of 1) visual inspection of photomasks used in the process, 2) in-line visual inspections, on a sample basis, of wafers in process, and 3) a heavy emphasis on end-of-line test data for initial electrical parameters. Certain mechanical properties such as solderability and various reliability parameters were controlled via lot sample tests. As noted in the paper by Stapper *et al.* [4], corrective action to process deviations can take a substantial amount of time. Because of the time lag in feedback response and the volume of product in process, one can expect relatively high average defect levels to occur when process deviations exist compared to control methods with inherently shorter time delays. Improving the response time in the discovery of process deviations became a prime factor in the reduction of defect levels.

Techniques such as the use of test sites, both on product wafers and on special or monitor wafers, evolved over a period of time. These test sites were essentially used to monitor key parameters controlling not only yields but also

the quality and reliability levels of finished product. A major improvement in process control of FET memory devices resulted when these test sites were introduced in the unused or kerf areas of the wafers and electrical parameters were measured after the first metallization process step. This test point has become known in IBM as *PAP* (post-aluminum probe) and has evolved as a key control point for the quality and reliability of integrated circuit devices. As density increased, additional schemes were introduced, such as control wafers, PAP site failure analysis, and cell failure mapping to achieve greater defect control within the fabrication process. These methods have been reviewed elsewhere in this issue [4].

SAMOS technology Silicon and metal oxide semiconductor (SAMOS) [5] technology represented the IBM entry into VLSI FET memory technology with the announcement of 64K-bit RAM devices in October, 1978. It also represented a stage in the 15-year evolution of monolithic memory manufacturing from a 16-bit bipolar device in 1965. In effect, the memory capacity of individual silicon chips has doubled each year during this time. As density has increased, device defect levels and reliability performance have improved by orders of magnitude (Fig. 1). But this improvement was not accidental. With the development and implementation of the SAMOS technology, reliability became of fundamental importance in the overall device design. This technology required consideration of the following major failure modes affecting the reliability of the device:

1. Dielectric charge stability under positive gate stress. Instability of the threshold voltage, V_T , in FET device structures usually occurs because of mobile charges created by ionic impurities resulting from an in-process alkali contaminating the dielectric. To ensure threshold stability under positive gate bias, a nitride process was developed and implemented [6] to stabilize the device structure.
2. Channel and substrate hot-electron injection into the dielectric under voltage stress—*i.e.*, hot-electron effects. Threshold voltage shifts under gate bias can also occur under conditions of high conduction when both gate and drain electrons encounter lattice scattering and are subsequently injected over the potential barrier at the oxide-silicon interface, the nitride gate acting as a trap. The effect of these channel hot electrons is an increase in V_T with time as the electrons accumulate in the gate isolator. A channel hot-electron degradation model was developed to predict threshold shift over time as a function of device design and insulator characteristics [7]. This model was applied to minimize the threshold shift by optimizing the depth of the drain junctions and doping levels.

An additional hot-electron effect, known as substrate hot electrons, is also generated by leakage current creat-

ing a leakage-induced threshold shift (LITS). These are electrons generated within the depletion region of a conducting gate. Injection occurs as the field near the semiconductor surface creates an electron energy comparable to the oxide-silicon potential barrier. The LITS effect is controlled by minimizing the oxide-silicon fields through the use of a high-resistivity p substrate and by controlling the ion implant doping.

3. Parasitic leakage from source to drain, or "sidewalk" leakage. This leakage mechanism can occur through the creation of a parasitic channel between source and drain after a positive gate stress. The horizontal charge movement that results is known as *sidewalk* and is caused by ionic contaminants near the edge of the metal gate in conjunction with recessed areas having entrapped contamination. This effect was eliminated by using a metal-oxide-nitride-oxide-silicon (MONOS) gate structure which leaves the oxidized nitride intact (Fig. 2).

These reliability mechanisms are primarily controlled in production through in-line monitoring of characteristics of metal-oxide-silicon or MOS capacitor wafers which are processed together with product wafers through the arsenic drive-in step of the process. Additional controls are exercised through the use of test site capacitors designed into the product wafer. Pulsed capacitance measurements [8] are used to predict wafer doping levels and flatband voltage characteristics earlier in the process.

As noted previously, various device parameters are controlled through electrical probe-type measurements on key test sites after aluminum evaporation. These so-called PAP measurements have been supplemented by visual inspection of defects resulting from the photolithographic process known as PLY or photo-limited yield inspection. Originally developed through the work of Dennard [9], PLY data have been applied to a photo-limited defect reliability model. Employing this model, one can predict reliability performance caused by this type of defect from visual inspection data.

In general, the IBM SAMOS production lines are controlled for all known functional stress failure mechanisms. A composite mathematical model of these mechanisms is used to predict product reliability from in-line indicators on a continuous basis. This approach to quality control is described in the following section.

Quality approaches in SAMOS

The manufacturing process for the SAMOS technology consists of a complex sequence of so-called hot-process, photolithographic, and metallization steps [5]. The strategy in the manufacturing of this technology is to build quality into the product through a system of statistical control

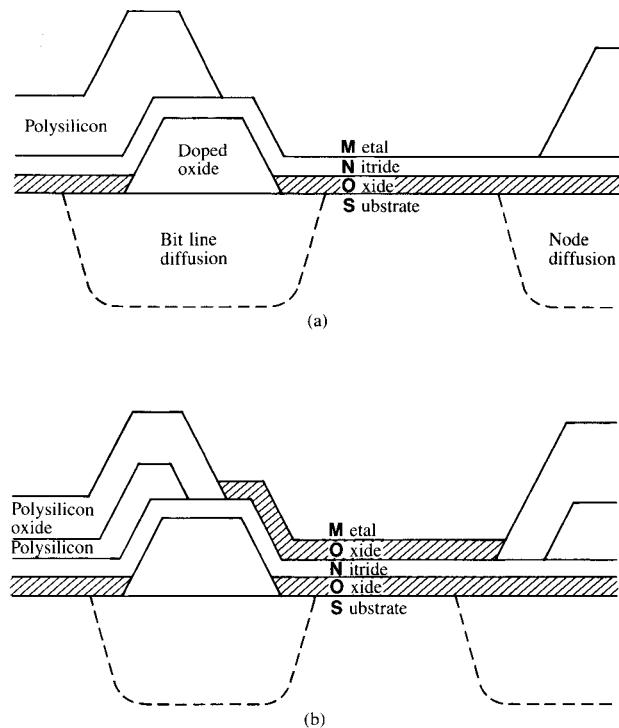


Figure 2 (a) An ordinary metal-nitride-oxide gate structure. (b) The improved MONOS gate structure showing the additional oxide isolation layer.

schemes on a series of yield- and reliability-related parameters which affect product performance.

Process control schemes, for the purpose of achieving devices that meet the quality and reliability objectives specified for the product at the end of the line, have been developed for properties such as foreign material levels, oxide thickness distributions, photo-defect levels, probed electrical measurements, and others. The key control techniques essential to SAMOS quality are in-line visual control, pinhole defect control, reliability stress tests, and feed-forward modeling.

In-line visual control At each photolithographic step in the fabrication process, PLY inspections are performed in which operators inspect devices by microscope for various types of visual defects affecting yield and reliability. This PLY information is summarized for each step or level in the photolithographic process by product type. From the data, a PLY value is calculated for each process step as described in the Appendix. This value is incorporated into a functional yield model along with other yield-limiting values to project functional yields.

For SAMOS, eleven types of visual defects have been identified which can cause reliability failures in device

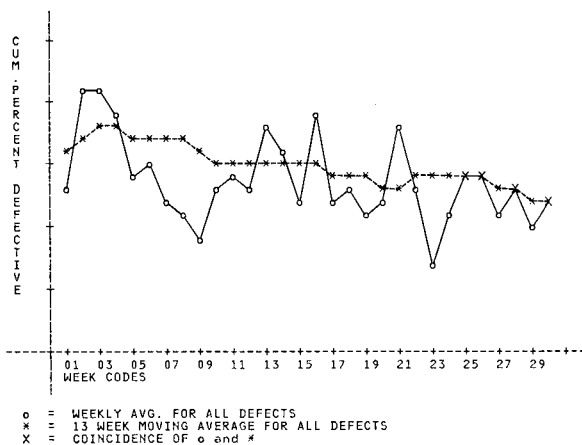


Figure 3 Cumulative visual impact for reliability-type defects. This plot displays

$$\sum_{i=1}^{11} p_i \times F_i,$$

where p_i = process average for visual defect i , and F_i = visual field factor for defect i .

performance. These were determined through detailed electrical and physical analysis of device failures occurring during accelerated forms of reliability stress testing. Foreign material in the gate region, reduced gate dimensions caused by the etching photoresist operations, defects in the oxide, and misalignment-caused defects have been found to be the major problems affecting reliability.

In order to have an indication of the possible reliability impact of shifts in the process average for these eleven defects, field factors are derived for each defect type. A field factor is a weighting factor which, when multiplied by the process average, gives an estimate of the reliability contribution caused by the specific defect. These field factors were derived from analysis data of failures from accelerated life testing, the aggregate reliability results, and the defect process average for this vintage of product. The field factor γ for the j th defect type is calculated using the relationship

$$\gamma_j = \frac{\delta \rho_j}{\mu_j}, \quad (1)$$

where δ represents aggregate reliability results, ρ_j is the proportion of the j th type of defect to the total found in the analysis, and μ_j is the process average for the j th defect. The total reliability impact caused by photo-type defects is estimated weekly by product type. To do this for a particular vintage of production, the product of $\gamma_j \mu_j$ is calculated and summed for the total set of defect types and compared with engineering specifications. This is a part of the feed-forward modeling to be discussed subsequently. Figure 3 shows a

sample plot of reliability-type visual defects on reliability performance for a 64K-bit RAM chip product, showing both the weekly average for all defects and a thirteen-week moving average.

Statistical control limits exist on each of the reliability defect types. Standard sampling techniques are used for visual inspection of each defect type that can be observed at specific photo levels. Each type of defect has a target for its process average (that is, the percentage of inspected chips with that reliability defect). Each defect type is monitored as to its level of occurrence to assess whether its process average is in control. When the process average of a particular reliability defect exceeds the upper control limit established at a 99% confidence level for a group of lots, corrective action is immediately taken on the assignable cause. Typical actions taken are removal of a particular tool in the process sector causing the problem and performing the appropriate adjustment or modification so that performance within the capability of the process can be maintained.

Pinhole defect control In FET technology, pinhole defect failures have been identified as one of the predominant failure mechanisms under both stress and application conditions. For SAMOS, the mechanism is vertical leakage from a diffused region in the substrate through the gate oxide and silicon nitride layers to the polysilicon field shield. This is known as a silicon-nitride-oxide-silicon or SNOS failure mode. The predominant cause of SNOS failures is submicron-sized material protruding through these layers. To control the line for this failure mechanism, electrical tests and various visual inspections for foreign material have been implemented at strategic points in the process.

A number of different particle inspection techniques have been developed for both monitor and product wafers. The techniques involve oblique light- and dark-field inspection of product wafers as well as high-magnification photographs of monitor wafers. These inspection techniques have been implemented, data gathered, and control limits derived. Although they are used throughout the line, these controls are especially important at the gate oxidation and silicon nitride/polysilicon deposition sectors for controlling and decreasing the source of SNOS failure.

In addition to controlling the particle levels of the line, there are also electrical tests for controlling the SNOS mechanism. The first of these occurs after the gate oxidation and silicon nitride and polysilicon depositions have been completed and the FET gate window opened. Both nitride deposition tool monitors and some product samples are tested at this point. The monitor wafer consists of a bare silicon wafer with a deposition of gate oxide and nitride layers. For this process, gate oxide is thermally grown while nitride is

deposited. A series of aluminum dots used as probe sites are evaporated on the nitride. In the SNOS test, a stress voltage is applied and the induced dielectric leakage current is measured and compared to the failure criteria. From this data, a quality index (QI) value is calculated which represents the percentage of sites with leakage currents below the failure criteria.

The QI of a monitor wafer is essentially a statistical description of the leakage profile of the wafer. Individual QIs are computed to form a daily index for each evaporation tool. Control limits on nitride deposition tools have been derived to preclude degradation from demonstrated process capability. Each tool, then, is required to meet a minimum index of performance in terms of leakage; otherwise, it is removed from production for corrective action. This test is also made on product wafers by actually probing the polysilicon field shield and applying a similar test sequence. Five wafers of a product lot are tested on a skip lot frequency, and the line is monitored against a lot QI average requirement of 95%.

Reliability stress tests After the PAP test, in-line reliability stress tests are done on a random sample of product wafers daily to control the line for two of the previously described reliability mechanisms, namely threshold voltage stability, ΔV_T , and sidewalk, S_w .

The ΔV_T parameter is a negative shift in threshold voltage caused by mobile charge contamination in the dielectric as mentioned before. Tests to control this parameter are done on a kerf site and consist of measuring the gate voltage necessary for a specified drain current, applying a voltage/temperature stress to drive mobile charges toward the substrate, and remeasuring the gate voltage. Corrective action is required if the averages for threshold shift exceed a specified value on the basis of ten-wafer samples.

The sidewalk stress is done on product kerf sites in conjunction with the ΔV_T test. This parameter measures the creation of a parasitic n-channel FET under voltage and temperature stress. Leakage current criteria are set for the particular stress levels in conjunction with the reliability model applicable to this failure mode. Failure limits are generally set in the picoampere range for this type of VLSI device.

Feed-forward reliability modeling

Failure analysis on devices subjected to extended reliability tests has resulted in identification of major failure modes, namely SNOS, visual defects, mobile charge failures, and polysilicon oxide defects. In the same way that field factors were developed for the eleven in-line visual defects to predict the impact on reliability performance, reliability factors have also been derived for other in-line tests to project the impact

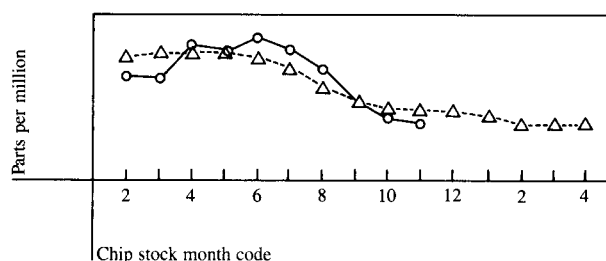


Figure 4 Tracking actual reliability performance (O) against projected performance (Δ) on a monthly basis.

on these failure modes. Field factors have been derived for the SNOS failure mode using the device-limited yield data from PAP measurements. Sidewalk yield data from PAP measurements are used to predict the mobile charge mode, while metal-to-polysilicon pinhole yield data at PAP are used to predict the polysilicon oxide failures. Monthly in-line defect indicators are used, in conjunction with field factor values for individual failure modes, to project the 100-hour reliability performance for any production vintage. Figure 4 is a plot of actual and projected reliability performance. The tracking permits management to anticipate future reliability trends, and to plan manufacturing burn-in requirements.

A so-called reliability *express* system has been developed to aid in the modeling of reliability performance from in-line indicators. With this system, product with known in-process characteristics is processed on a send-ahead or expedited basis for reliability stress tests. This allows for rapid assessment of the reliability performance of the production vintage in process and the taking of appropriate action when it is deemed necessary.

Control techniques

In addition to visual inspections and electrical tests for defects for which field factors have been developed, a large number of other parameters are measured electrically or optically and have associated control schemes. Some of these relate to yield performance, while the interaction of others may relate to the reliability performance of SAMOS. Measurement and control schemes for oxide thickness as well as fixed and mobile charge levels in the grown oxide have been implemented at each of the steps in the hot-process sector.

An example of some of the techniques employed to control the hot-process operation is gate oxidation. At this step, 100 wafers are oxidized at a time. Statistical analyses had indicated that there was a thickness ramp or profile across the boat during the process qualification period. As a result, thickness measurements on product and monitor wafers at extreme positions are required after each run. Also, during the qualification period, the arithmetic means for the gate

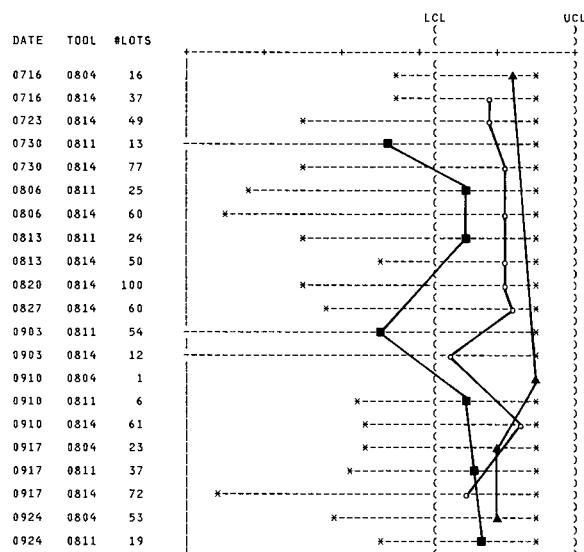


Figure 5 Trend chart for tool number 804 (▲), number 811 (■), and number 814 (○). Note that tool number 811 performed below the lower control limit (LCL) on 7/30 and 9/3.

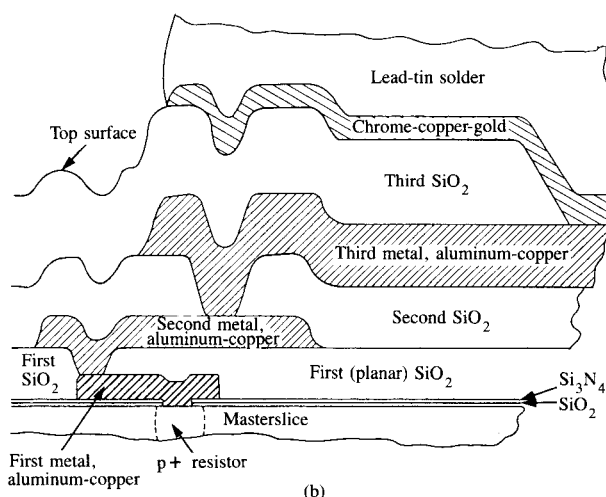
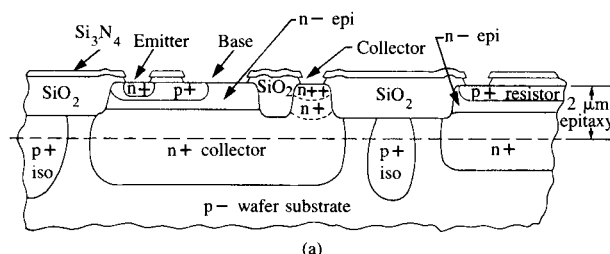


Figure 6 (a) Schematic of bipolar product masterslice cross-section. (b) Same product after terminal metal "personalization" in the p+ resistor area.

oxide thickness range for a wafer was determined and control limits established. Normally, five measurements per product wafer are made at pre-defined positions. Regression analyses were performed to establish limits for a single measurement which demonstrates at a 99% confidence level that virtually all sites on the wafer would be within specified limits for oxide thickness. If the single site measurement exceeds these statistically derived limits, additional measurements are made. When control limits on thickness are exceeded, product and tool action is taken in order to guarantee the specification for this parameter. Typical actions taken are nonconforming materials disposition reviews and removal of the deposition tool from production for maintenance.

Data system

Essential to maintaining quality control of the SAMOS product is a continuous and real-time assessment of the key parameters governing the quality and reliability of the product. Vast amounts of test measurement data and logistics data gathered through the key process operations are stored in a data base by individual monitor and product wafer serial identification numbers. These data can be retrieved with options for generating trend charts, control charts, histograms, correlations, and other statistical reports. For problem solving, the system allows one to recall PAP kerf defect data and organize the data in the sequence that the associated wafers went through at a previous point in the process. This is termed a *timeslide* of the data. An example of this is given in Fig. 5, where metal-polysilicon pinhole DLY data as measured at PAP are "timeslid" through polysilicon deposition by the particular tool used. This indicates a poorer performance for deposition tool number 811, which was subsequently shut down.

These examples illustrate the methods employed in the approach to quality manufacturing of SAMOS products. In addition to statistical controls for a stable process, this approach also means using sound statistical techniques for evaluating the impact of proposed process changes, as well as ensuring that the manufacturing of parts is done in accordance with documented procedures. This total system approach is, in part, responsible for the success of the SAMOS manufacturing line.

Other features also exist which aid in line control, such as a merge-code system for identifying a segment of production, and a flagging system to segregate product at the end of the line for dispositioning actions. The process profile management system to be described illustrates further applications of the manufacturing data base to quality control.

IBM's bipolar technology

The manufacturing process for IBM's high-performance bipolar logic product, as with IBM's FET technology, con-

sists of a sequence of hot-process, photolithographic, and metallization steps. The bipolar device process is divided into two basic processing sectors, namely masterslice and "personalization." The masterslice portion is a series of repeated diffusion-oxidation and photolithographic process steps where all the pn junctions, p and n resistors, transistors, and Schottky diodes are formed within the body of the silicon wafer. In personalization, the active and passive devices are metallurgically connected and insulated in a multi-level structure to form functional logic and array circuitry [10]. Figure 6 shows the basic bipolar structure of the masterslice before and after personalization.

Process control approach

With complex processes, achieving a viable quality control system requires a concerted effort between the quality and manufacturing process engineering people to define control points and measurement methods. The system for controlling device manufacture covers the following major areas:

- photolithography control,
- hot-process control,
- pn junction control.

Photolithography control The key factors in the photolithographic process are pattern fidelity and contamination control. These factors lead directly to defect detection by visual inspection at the various photolithographic stages. The visual inspection data are fed into a PLY model as a measure of process defect levels and a predictor of performance. This PLY control approach has been described previously and it applies equally to bipolar technology.

Hot-process control The diffusion process and its associated thermal oxidation step are so fundamental in determining device performance and yields that its characteristics are measured as soon as practicable in the process sectors rather than at end-of-line. Because of this need for process data to allow for timely corrections to the process, measurements are done on monitor wafers. These monitors are wafers without device patterns but with a well-defined background of impurity concentration to provide a contrast to the diffusants being controlled. The characteristic of control is the profile of the diffusants. This is typically done with a *spreading-resistance* probe technique whereby the specimen is beveled to allow for the probe heads to traverse the junction to determine the concentration of impurities at various points, as shown in Fig. 7(a). The profile as shown in Fig. 7(b) characterizes the diffusion process which in turn determines the device performance.

Because measurement of spreading resistance is sophisticated and time-consuming, it is done mostly in the laboratory environment. What is routinely performed in practice is a four-point probe measurement. This measurement provides

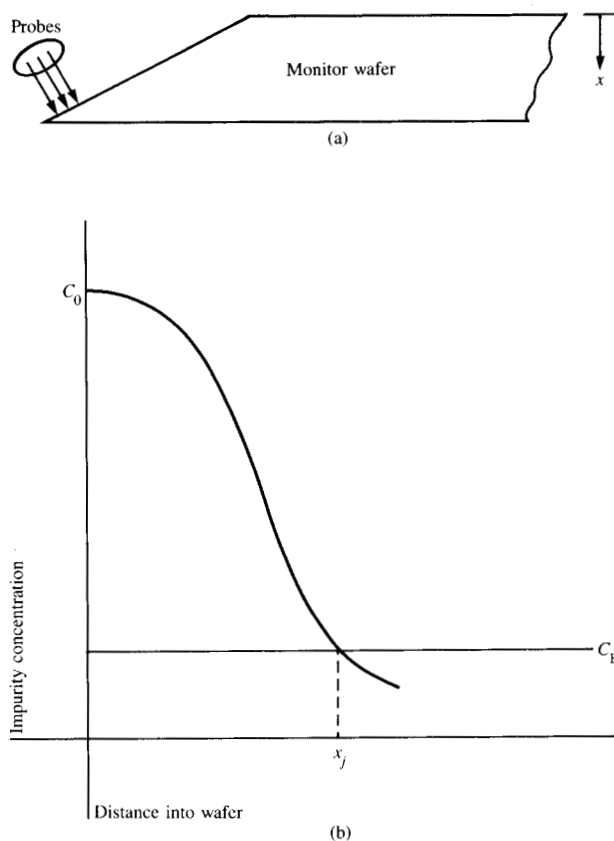


Figure 7 (a) Spreading resistance measurement over a beveled monitor wafer. (b) Impurity profile plot, where C_0 is the surface concentration, C_B is the background concentration, and x_j is the junction depth.

an average resistance, commonly known as sheet resistance R_s , which is related to the impurity profile through the well-known relationship

$$R_s = \frac{\bar{\rho}}{x_j} = \frac{1}{q \int_0^{x_j} \mu n(x) dx}, \quad (2)$$

where $\bar{\rho}$ is the average resistivity expressed in ohm-cm, μ is the mobility of the charge carriers, and $n(x)$ is the shape of the profile. Examples of statistical data on sheet resistance for the various oxide steps are shown in the process profile in Fig. 11, which appears later.

pn junction and device parametric control For control of junction and parametric properties, device characteristics are obtained through a series of in-line electrical measurements on specially designed structures, either in the kerf areas or on test sites. The key concern here is junction quality as defined by breakdown voltages and leakage current, and manifested by *pipes* and/or mobile charges. Many device parameters are measured and monitored in-process. The key parameters are isolation breakdown voltage BV_{ISO} , which monitors the integrity of device isolation; collector-to-base

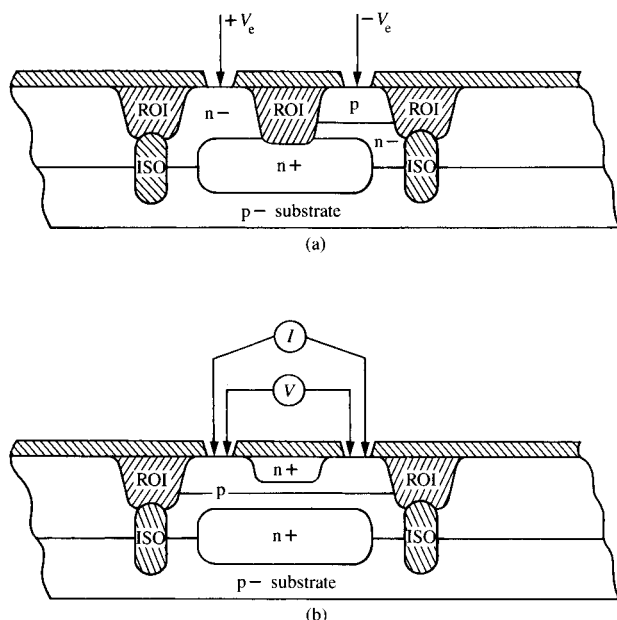


Figure 8 (a) BV_{CB} measurement, where ROI is recessed oxide isolation and ISO is isolation oxide. (b) R_{DB} measurement, also called four-point probe measurement.

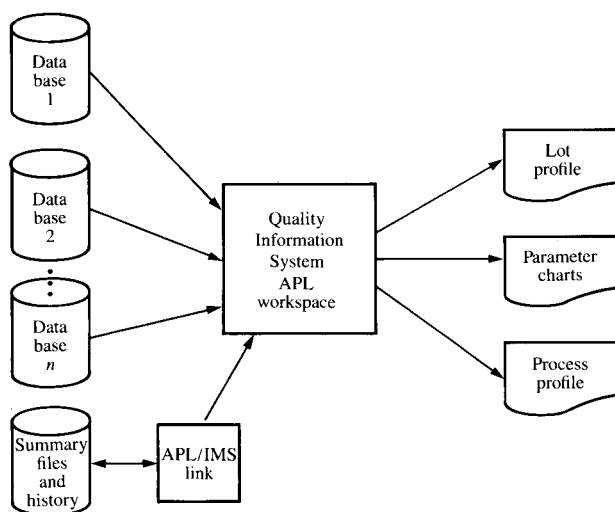


Figure 9 Quality Information System overview.

breakdown voltage BV_{CB} [see Fig. 8(a)]; and collector-to-emitter breakdown voltage BV_{CE} . BV_{CB} and BV_{CE} control pipe and epitaxy variables. Electrical measurements of these parameters are made at various process points at the completion of certain device junctions.

Also used as a process control device is the so-called dumbbell resistor. Its structure is the closest approximation to an actual transistor and it is a powerful tool in the diagnosis of transistor defect properties. As is well known, a

low resistance R_{DB} is an indication of insufficient emitter diffusion, too deep a base diffusion, or too high a boron concentration. Conversely, a high R_{DB} may be an indication of too deep an emitter diffusion, a shallow base diffusion, or a low boron concentration. Figure 8(b) illustrates this measurement.

Transistor gain, or β , is often taken as the figure of merit for bipolar devices. As is well known, careful control of base width is the primary consideration for keeping β within specification. Low values are often associated with wide base width. Conversely, high values are associated with narrow base width. Figure 6(a) provides a schematic of the bipolar transistor structure and also shows the criticality of base dimension control.

These parameters provide the bulk of the process control information for feedback to process engineers for corrections as well as feed-forward information to diagnostic engineers to predict product performance, namely yield, quality, and reliability.

Product reliability

As with the SAMOS technology, reliability defects and their causes have been identified for bipolar devices, and in-line monitors have been devised. This makes possible the direct measurement of product for reliability criteria as it is being manufactured. Device reliability is ensured in two principal ways: by controlling the process and by stress-testing various defect monitors and reliability stress vehicles. Defect and stress data are used in predicting the failure rate of each masterslice type periodically through use of reliability models. (An example of this type of model is the electromigration model which predicts the reliability of aluminum metallization as a function of cross-sectional area and current density [11, 12]). These models take into account various masterslice differences such as metal crossover areas, voltage levels, device densities, and current densities, and they relate empirically determined process defect types at specific levels to ultimate failure rates. Reliability performance feedback measurements at higher levels of assembly, such as circuit cards and shipped systems, make up the total reliability performance picture of the product.

Defect-induced failures are the largest contributors to reliability problems at the device level. Since the low failure-rate level associated with VLSI would require a prohibitive quantity of samples to be evaluated by traditional end-of-line stress methods, direct process monitoring and control is the only viable course for ensuring quality and reliability. In addition, defects that can cause these failures cannot be adequately monitored on the completed chip, and so direct process monitoring of defects is used to control these type failures. The major failure modes associated with the bipolar

device process are interlevel shorts and opens in metal lands. Test sites consisting of special metallization patterns are applied mid-process for test purposes and are used to measure interlevel shorts. This is done by processing test sites with regular product runs and then subjecting them to voltage stresses to induce failures. Metal land failures are monitored by product inspections that look for defects that contribute to reduced cross-sectional areas of metallization.

The impact on data analysis of monitoring numerous in-process parameters presents a significant challenge to the methods by which product quality and reliability are to be ensured. Methods had to be developed that could provide the total picture of product manufacture in terms that could be easily perceived with minimum inspection resources and in a meaningful timeframe. The system which was used to convert the vast amount of reliability and in-process measurement data into statistical summary information on the product/process status is described in the following sections of this paper.

Process profile

With sophisticated technologies and process complexity, translation of a collection of individual control charts into a meaningful assessment of the overall process is at best difficult. Specifications are more detailed and requirements for product quality are significantly more stringent. Further, not only is product manufacture more complex, but there is a proliferation in the number of product types involved. Other factors making analyses more difficult are the increased number of parameter interactions and more sophisticated tooling, which require closer attention to product performance and process quality assessments.

A system overview depicting data base access and output report capability is shown in Fig. 9. The system accesses existing manufacturing data bases, making it not only independent but easily tailored to function with any existing data base. The approach developed uses what is known as a *process profile*. Basically, a process profile takes the results of measurements at each process step for a specific time period and summarizes them in terms of mean and range values. Upper and lower control limits equal to three standard deviations are substituted for the range when performing process capability studies; but the range was found to be of most benefit for routine line monitoring.

Figure 10 shows the basic indicators of process performance that are obtained from a process profile. The $\times$ represents the position relative to specifications of the parameter mean value. The data spread of the parameter is shown by the solid line going through the $\times$ mean position. The number symbol (#) shows where an engineering-specified (ES) limit is placed. The mean, high, and low points are

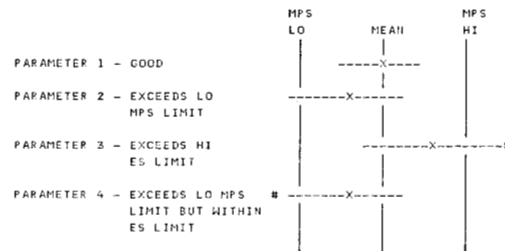


Figure 10 Process profile definitions. MPS stands for manufacturing process specifications; ES for engineering specifications; the number symbol (#) is defined in the text.

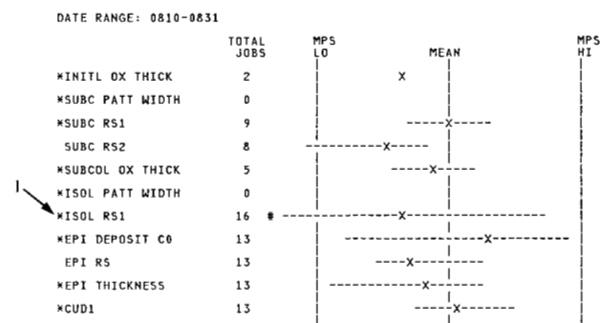


Figure 11 Typical quality information system masterslice process profile analysis results illustrating a potential problem.

normalized against manufacturing process specification (MPS) limits in relative rather than absolute values. This permits a direct comparison of all parameters in common terms.

An example of a process profile is shown in Fig. 11. The operations are arranged in order of process flow from top to bottom. This provides, at a glance, a picture of how the entire process is performing for a given period of time and which parameters need attention. It should be noted that this profile chart does not display a time sequence, but rather a time period. An asterisk appears as a flag preceding the parameter name if the parameter has an engineering specification associated with it. Parameters with engineering specifications also show a number symbol on their plot line at the location of the specification if they exceed a process limit, so that the severity of failure can be better assessed. Additional information provided is the number of jobs started into the line during the previous four weeks, the date range that it represents, and the masterslice or product involved.

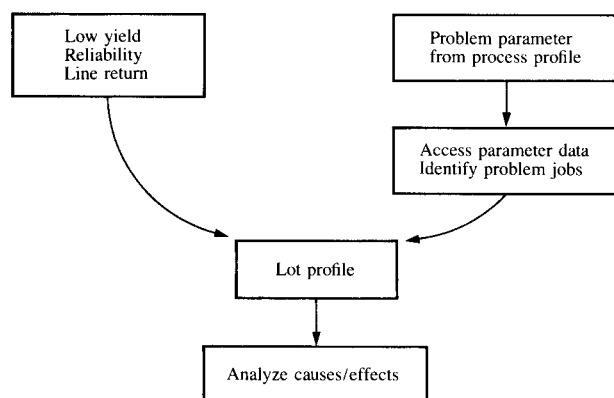


Figure 12 Typical problem trace routes.

A lot profile is similar in concept to the process profile, with the difference that the data points represent data for a specific job in question. Again, the data points are arranged in the order of processing. This allows immediate capture and display of all relevant process data for a single group of product on a single sheet, thereby facilitating rapid cause-and-effect analysis.

Both process and lot profiles are particularly suitable for use as problem analysis tools in semiconductor manufacturing. They provide all normalized control limits and parameters of interest for the entire semiconductor device process in a single graphical display. An important advantage of doing this is that parameter interactions can be assessed almost immediately once the chart is constructed. Since the process profile is based on cumulative data for a group of jobs in a given time period, it becomes an indication of the general state of affairs in the line, or "health of the line," for that period of time. It highlights suspect or problem areas by displaying the actual distribution against the control limits.

Figure 12 depicts how problems are traced back for analysis. The left-hand side of the figure demonstrates how feedback from the end-of-line, from final test, or from a downstream application is used to activate an analysis. For example, a specific job or production run may be identified as not performing properly, either at final test (low yield), in subsequent levels of assembly, or after a number of hours of application use. Since the job number is known, one can proceed directly to the lot profile and investigate all parameter values and their relationships to specification limits, in order to gain an understanding of the causes for failure. The right side of the figure shows how a problem parameter would be highlighted by the process profile and traced back for analysis. Such problem parameters will reveal themselves as the distribution extends beyond the control limits.

Once the problem parameters are identified, a *parameter plot* for the next level of analysis is obtained. This plot is a conventional control chart that plots the distribution of data from all jobs that were included in the process profile, but taking one parameter at a time. Since job numbers are associated with the corresponding data, one can easily pick out those jobs that are responsible for that parameter being out of limits in the process profile. In general, most of the reasons relating to the problem parameter are known at this stage. However, the system is designed so that an even finer level of analysis is possible by using the lot profile option. Lot profile analysis is used for investigating causal relationships among the parameters. Since the lot profile contains data from one given job, any one-to-one relationship is more likely to be shown here than on the other charts. By utilizing these profiles, one can identify problem areas and problem jobs from a massive array of complex data base elements.

An example of the use of the process profile system as a monitoring and corrective action tool is shown in Figs. 13 and 14. Figure 13 shows a process profile for the time period 9/8 to 9/14. Picking up on Problem 1, where the subcollector oxide thickness parameter is shown to be failing the upper engineering specification limit, a parameter plot as shown in Fig. 14 was obtained. The problem was then traced to a specific job (number 3347) that exceeded the specification limit. Further investigation by the quality department revealed that the job had been sent to the next operation without corrective action. This investigation triggered a manufacturing procedure to prevent recurrence of this error. The system thus provides a powerful tool to ensure conformance of a process to prescribed requirements and correction where requirements are inadequately specified.

Reliability monitoring

In addition to the process profile that is used to monitor the overall process, a system has been established to monitor specific reliability parameters at both the detailed and summary levels. As noted before, the product is inspected and sample tested for characteristics that can influence reliability at various stages of manufacture. This information is entered into a general manufacturing data base. Product is monitored on individual-job bases and through weekly summaries for manufacturing line control. In addition, the data from these inspections are summarized on a monthly basis and these summary data points are stored in a reliability data base. These data are retained over the anticipated life of the product so that analyses of field failures can be performed if the need should ever arise. These summary data are primarily used to perform monthly and six-month cumulative defect level analyses against defined reliability specifications. All reliability requirements for the entire device line product mix are included in this data base so that a total system approach to product monitoring is facilitated. Evi-

dence of trends that could potentially degrade device performance is fed back into the process for correction of the cause. The system provides a comprehensive measurement of the reliability level at which the total process is operating at any given point in time.

The system also provides for exception trend reporting. For this reporting option, the entire data base is interrogated and trends are searched. Parameters that show definite trends are displayed. This capability provides management with exception information on specific process areas and device parameters requiring action.

Summary

The approaches used to control quality and reliability of two VLSI technologies currently in production at IBM have been described. We do not see these methods as necessarily the final solution, but rather as the starting point from which to meet the challenges of the technology evolution, both now and in the future.

Appendix

The PLY equation employs a series of λ values pertaining to the probabilities that chip failures occur for defect types 1 through k . These values are derived from characterization data after PAP. If, for example, there are k different yield defects inspected for at a certain photo level, the PLY of n chips is

$$PLY = (100) \sum_{i=1}^n \frac{(1 - \lambda_1)^{x_{i1}} (1 - \lambda_2)^{x_{i2}} \cdots (1 - \lambda_k)^{x_{ik}}}{n}, \quad (3)$$

where x_{ik} = the number of type k defects on the i th chip.

Acknowledgments

Sincere appreciation must be expressed for the efforts of a large number of colleagues at the IBM facilities in Essex Junction, Vermont, and in East Fishkill, New York, who over many years evolved the quality control, inspection, and measurement systems and the data bases, to bring them to the present state of the art which we have tried to summarize in this article. In addition, a number of people have made helpful suggestions on this manuscript. In particular, the assistance of C. Behr, F. Youton, D. Begin, C. Nierenberg, and B. Slade is gratefully acknowledged.

References

1. W. A. Shewhart, *Economic Control of Manufactured Product*, D. Van Nostrand Co., Inc., Princeton, NJ, 1931.
2. W. E. Deming, *Some Theory of Sampling*, John Wiley & Sons, Inc., New York, 1950.
3. E. M. Davis, W. E. Harding, R. S. Schwartz, and J. J. Corning, "Solid Logic Technology: Versatile, High-Performance Microelectronics," *IBM J. Res. Develop.* **8**, 102 (1964).
4. C. H. Stapper, P. P. Castrucci, R. A. Maeder, W. E. Rowe, and R. A. Verhelst, "Evolution and Accomplishments of VLSI Yield Management at IBM," *IBM J. Res. Develop.* **26**, 532 (1982, this issue).
5. Richard A. Larsen, "A Silicon and Aluminum Dynamic Memory Technology," *IBM J. Res. Develop.* **24**, 268 (1980).
6. P. K. Chaudhari, J. M. Franz, and C. P. Acker, "Electrical Properties of Vapor Deposited Silicon Nitride Films Measured in Strong Electric Fields," *J. Electrochem. Soc.* **120**, 991 (1973).
7. R. R. Troutman and A. G. Fortino, "Simple Model for Threshold Voltage in a Short Channel IGFET," *IEEE Trans. Electron Devices* **ED-24**, 1266 (1977).

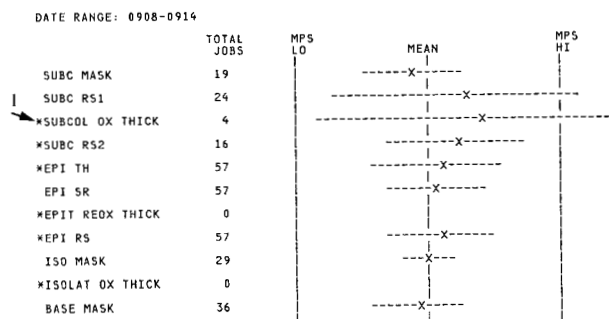


Figure 13 Using process profile analysis to follow up on problems. Here Problem 1 shows subcollector oxide thickness exceeding the high process specifications and just at the engineering specifications.

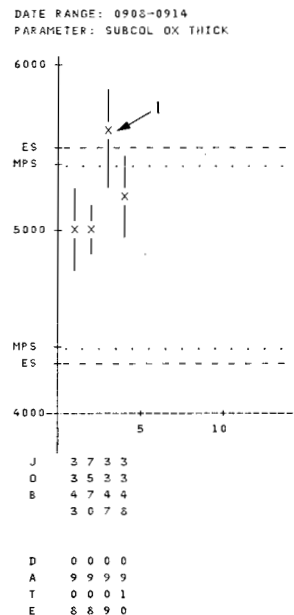


Figure 14 Subcollector oxide thickness parameter plot for Problem 1 illustrated in Fig. 13. Job number 3347 is found to exceed engineering specifications (ES).

8. G. E. Schmid, "Pulsed CV System for Ion Implantation," *Nuclear Instr. Methods* **189**, 219 (1981).
9. R. H. Dennard, "Cost Study for Integrated Circuits with Many Logical Decisions per Chip," *Research Report RC-1552*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, 1966.
10. H. W. Curtis, "Integrated Circuit Design, Production, and Packaging for System/38," *Electronics*, McGraw-Hill Book Co., Inc., New York, 1979.
11. D. S. Chhabra and N. G. Ainslie, "Open Circuit Failures in Thin Film Conductors," *Technical Report TR22.419*, IBM Components Division, East Fishkill, NY, 1967; presented at the Electrochemical Society Spring Meeting, Dallas, TX, May 7-12, 1967.
12. F. M. d'Heurle, N. G. Ainslie, G. Gangulee, and M. C. Shine, "Activation Energy for Electromigration Failures in Aluminum Films Containing Copper," *J. Vac. Sci. Technol.* **9**, 289 (1972).

Received September 17, 1981; revised March 23, 1982

Richard T. Curtis *IBM General Technology Division, East Fishkill facility, Hopewell Junction, New York 12533.* Mr. Curtis joined IBM in 1952 as a technician at the Poughkeepsie, New York, plant. He worked for five years on calibration and maintenance of production test equipment. In 1957, he joined the Quality Engineering Department and was responsible for the quality system in the ferrite memory core manufacturing and test areas. He has worked extensively on the application of data processing methods to the field of quality control. Mr. Curtis transferred to East Fishkill in 1967. He is currently an advisory systems analyst in the Quality Data Systems Department, where he is responsible for development and implementation of data processing applications for the quality function. Mr. Curtis is a senior member of the American Society for Quality Control.

Joseph K. Ho *IBM General Technology Division, East Fishkill facility, Hopewell Junction, New York 12533.* Mr. Ho graduated from the University of Illinois, Urbana, in 1967 with a B.S. in physics and obtained his M.S. in physics from Vassar College, Poughkeepsie, New York, in 1976. His research interests are in the areas of statistical interpretation of quantum mechanical measurements. He joined IBM's Components Division in 1967, working on thick film engineering and resistor trimming. In 1971 he joined device quality engineering, primarily in the area of epitaxial growth, oxidation, and thermal diffusion. He has worked in the areas of

optical and electrical characterization of semiconductor materials and the applications of such measurements as monitoring vehicles in the device manufacturing areas. Mr. Ho's current interests include the use of automated data systems for real-time on-line process controls.

Jeffrey G. Koens *IBM General Technology Division, East Fishkill facility, Hopewell Junction, New York 12533.* Mr. Koens is a development engineer currently responsible for managing the development and coordination of East Fishkill's manufacturing planning systems. He joined IBM in East Fishkill in 1969 as an associate engineer and has held a variety of management and technical positions in the development, manufacturing, and quality engineering areas. Mr. Koens received a B.S. in physics in 1968 and a bachelor of electrical engineering degree in 1969 from the City University of New York.

Eugene H. Melan *IBM Data Systems Division, P. O. Box 390, Poughkeepsie, New York 12602.* Mr. Melan is a senior engineer on the technical staff of the Data Systems Division and is currently responsible for the division quality excellence program. He joined IBM in Poughkeepsie in 1954 as a junior engineer working on the development of the magnetic drum unit of the IBM Model 702. Subsequent assignments in magnetic technologies included development and characterization of ferrite and film devices for memory and switching applications. He has held a variety of management positions in development and manufacturing since 1960, including device analysis and reliability, advanced module engineering, current product assurance, supplier technical assurance, and world trade technical programs. He received a B.S. in mathematics in 1953, an M.S. in physics in 1954 from New York University, and an M.S. in industrial administration from Union College, Schenectady, New York, in 1972. Mr. Melan is a member of the American Society for Quality Control and Sigma Xi; he is an ASQE certified quality engineer.

Gary A. Snyder *IBM General Technology Division, Burlington facility, Essex Junction, Vermont 05452.* Mr. Snyder is a senior associate engineer working in SAMOS reliability and characterization quality engineering. He is responsible for characterizing line problems and estimating future reliability performance via models utilizing in-line data of key reliability-related parameters. Mr. Snyder joined IBM in 1978 at Burlington as a quality engineer responsible for SAMOS hot process controls. Prior to joining IBM, he taught at John F. Kennedy High School, Warren, Ohio, and at Boardman High School, Youngstown, Ohio. He received his B.S. in mathematics from Gannon College, Erie, Pennsylvania.